Vitesse de convergence de l'algorithme de Metropolis–Hastings

L'objectif de ce devoir est d'étudier le temps de mélange d'une version de l'algorithme de Metropolis, qui permet de simuler une loi π sur un espace d'états $\mathfrak{X}$. Dans tout ce qui suit, on fixe un ensemble *fini* $\mathfrak{X}$ et une mesure de probabilité π sur cet ensemble : $\sum_{x\in\mathfrak{X}} \pi(x) = 1$. Quitte à réduire la taille de $\mathfrak{X}$, on peut supposer $\pi(x) > 0$ pour tout $x \in \mathfrak{X}$. On fixe également une matrice de transition $Q = (Q(x,y))_{x,y\in\mathfrak{X}}$ sur $\mathfrak{X}$, qu'on suppose irréductible. Cette matrice Q n'a *a priori* aucun rapport avec π (en particulier, la mesure invariante de Q n'est pas supposée être égale à π). On impose en revanche l'hypothèse simplificatrice suivante : si $Q(x,y) > 0$ pour une paire d'états (x,y), alors on a aussi $Q(y,x) > 0$.

1 Forme générale de l'algorithme de Metropolis–Hastings

Soit $(A(x,y))_{x\neq y}$ une collection de nombres réels compris entre 0 et 1, indexée par les paires d'états (x,y) de $\mathfrak{X}$ avec $x \neq y$. On appelle *matrice de Metropolis–Hastings de matrice génératrice* Q *et de probabilités d'acceptation* A la matrice stochastique définie par :

$$P(x,y) = \begin{cases} A(x,y)\,Q(x,y) & \text{si } x \neq y, \\ Q(x,x) + \sum_{x\neq y}(1-A(x,y))\,Q(x,y) & \text{si } x = y. \end{cases}$$

1. Vérifier que P est bien une matrice stochastique.

Les transitions de la chaîne de Markov $(X_n)_{n\in\mathbb{N}}$ associée à la matrice P se réalisent comme suit : si l'on est au temps n à l'état $X_n = x$, on choisit un état-candidat y pour la transition $X_n \to X_{n+1}$ suivant la loi de transition $Q(x,\cdot)$. Si cet état y est différent de x, alors :

— conditionnellement au fait que l'état-candidat soit y, on réalise effectivement la transition $x \to y$ ($X_{n+1} = y$) avec probabilité $A(x,y)$;

— avec probabilité conditionnelle $1 - A(x,y)$, on reste en l'état x.

L'algorithme de Metropolis–Hastings correspond au choix de coefficients d'acceptation $A(x,y)$ tels que la mesure invariante de P soit la loi π fixée initialement. Soit $S = (S(x,y))_{x\neq y}$ une famille symétrique ($S(x,y) = S(y,x)$ pour tout couple (x,y) d'états distincts), et

$$T(x,y) = \begin{cases} \frac{\pi(x)\,Q(x,y)}{\pi(y)\,Q(y,x)} & \text{si } Q(x,y) > 0, \\ 0 & \text{si } Q(x,y) = 0. \end{cases}$$

On pose alors $A(x,y) = \frac{S(x,y)}{1+T(x,y)}$.

2. On suppose que $0 \leq S(x,y) \leq 1 + \min(T(x,y), T(y,x))$ pour tout couple (x,y). Montrer que la matrice de Metropolis–Hastings P de matrice génératrice Q et de probabilités d'acceptation A admet pour mesure réversible (et donc invariante) π.

3. Donner une condition suffisante simple sur la matrice S pour que la matrice stochastique P soit irréductible et apériodique. En déduire que dès que $\mathrm{card}\,(\mathfrak{X}) \geq 2$, il existe une infinité de matrices stochastiques irréductibles apériodiques P pour lesquelles π est la mesure réversible.

Une matrice stochastique P privilégiée pour l'algorithme de Metropolis correspond au choix extrémal pour les coefficients $S(x,y)$: dans tout ce qui suit, on suppose que

$$S(x,y) = 1 + \min(T(x,y), T(y,x)) \quad \text{pour tout couple } x \neq y.$$

4. Montrer que sous ces hypothèses, les coefficients d'acceptation $A(x,y)$ s'écrivent

$$A(x,y) = \begin{cases} \min\left(1, \frac{\pi(y)\,Q(y,x)}{\pi(x)\,Q(x,y)}\right) & \text{si } Q(x,y) > 0, \\ 1 & \text{sinon.} \end{cases}$$

Que dire de la matrice P si π est réversible pour Q ? Montrer dans tous les cas que si Q est irréductible apériodique, alors P est irréductible apériodique de mesure réversible π.

2 Vitesse de convergence et forme de Dirichlet

On rappelle l'inégalité suivante : si P matrice stochastique irréductible admettant une mesure réversible π a pour valeurs propres $1 = \lambda_1 > \lambda_2 \geq \cdots \geq \lambda_N$, alors la chaîne de Markov $(X_n)_{n\in\mathbb{N}}$ associée à P vérifie pour tout temps n

$$d_{\mathrm{TV}}(\pi_n, \pi) \leq C\,\mathrm{e}^{-\rho n}, \quad \text{avec } C = \frac{1}{2}\sqrt{\max_{x\in\mathfrak{X}}\left(\frac{1}{\pi(x)} - 1\right)},$$

où $\rho = \min(1 - |\lambda_2|, 1 - |\lambda_N|)$ est le trou spectral de P, et $\pi_n = \pi_0\,P^n$ est la loi marginale de X_n. L'objectif de cette section est d'obtenir des estimées simples de ce trou spectral. On va introduire à cet effet la *forme de Dirichlet* d'une paire (matrice stochastique P, mesure réversible π).

5. On fixe une matrice P de loi réversible π (par exemple la matrice définie dans la section précédente par la méthode de Metropolis–Hastings). On définit l'énergie de Dirichlet d'une fonction $f : \mathfrak{X} \to \mathbb{R}$ par la formule suivante :

$$\mathscr{E}(f) = \langle (I-P)f \mid f \rangle_\pi = \sum_{x\in\mathfrak{X}} \pi(x)\,((I-P)f)(x)\,f(x),$$

où I est la matrice identité et $I - P$ agit sur les fonctions à gauche (les fonctions sont vues comme des vecteurs colonnes). Montrer que l'énergie de Dirichlet se réécrit sous la forme

$$\mathscr{E}(f) = \frac{1}{2}\sum_{x,y\in\mathfrak{X}} \pi(x)\,P(x,y)\,(f(x) - f(y))^2,$$

et donc qu'elle est positive. Si f est une fonction arbitraire et c est une constante, quelles sont les relations entre les énergies $\mathscr{E}(f)$ et $\mathscr{E}(f+c)$?

6. On considère une base de diagonalisation $(f_1, f_2, \ldots, f_N)$ de la matrice P correspondant aux valeurs propres ordonnées $1 = \lambda_1 > \lambda_2 \geq \cdots \geq \lambda_N$. Comme π est réversible par rapport à P, on peut supposer que cette base est orthonormée dans $\mathscr{L}^2(\mathfrak{X}, \pi)$. Montrer qu'on peut prendre pour f_1 la fonction constante égale à 1. Montrer aussi que toutes les autres fonctions $f_2, \ldots, f_N$ ont moyenne nulle sous π : $\pi(f_i) = \sum_{x\in\mathfrak{X}} \pi(x)\,f_i(x) = 0$ si $i \geq 2$.

7. La variance d'une fonction f sous la mesure π est donnée par

$$\mathrm{Var}_\pi(f) = \sum_{x \in \mathfrak{X}} \pi(x)\,(f(x) - \pi(f))^2.$$

Montrer que $\mathrm{Var}_\pi(f) = 0$ si et seulement si f est une fonction constante. Montrer ensuite la caractérisation suivante de la seconde valeur propre λ_2 de P :

$$1 - \lambda_2 = \inf \left\{ \frac{\mathscr{E}(f)}{\mathrm{Var}_\pi(f)} \mid f : \mathfrak{X} \to \mathbb{R},\ f \text{ fonction non constante} \right\}.$$

On pourra décomposer une fonction f sur la base $(f_1, \ldots, f_N)$ de la question précédente.

8. Montrer aussi que pour toute valeur propre λ_i de P, il existe $x \in \mathfrak{X}$ tel que

$$|\lambda_i - P(x,x)| \leq 1 - P(x,x).$$

On pourra considérer x tel que $|f_i(x)|$ soit maximal. En déduire que $1 + \lambda_N \geq 2 \min_{x \in \mathfrak{X}} P(x,x)$.

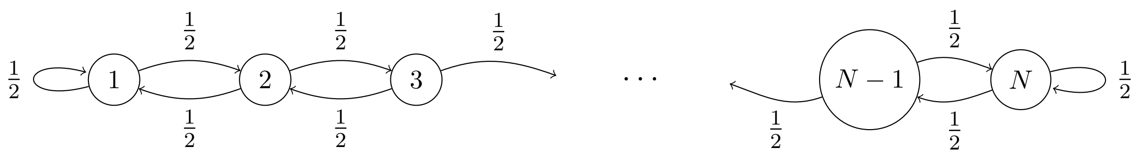
3 Simulation d'une loi à décroissance exponentielle

Dans cette dernière section, on considère l'espace des états $\mathfrak{X} = [\![1, N]\!]$, et une mesure de probabilité sur $\mathfrak{X}$ qui s'écrit sous la forme

$$\pi(x) = \frac{1}{Z_a}\, a^{h(x)},$$

où $a \in (0,1)$, $h : [\![1, N]\!] \to \mathbb{R}$ est une fonction telle que $h(x+1) - h(x) \geq 1$ pour tout $x \in [\![1, N-1]\!]$, et Z_a est la constante de normalisation telle que $Z_a = \sum_{x=1}^N a^{h(x)}$. On a donc une loi à décroissance au moins exponentielle, concentrée sur les petits entiers de $[\![1, N]\!]$.

9. On considère la matrice de transition Q associée à la marche aléatoire symétrique sur $[\![1, N]\!]$, avec probabilités de rebond $\frac{1}{2}$ aux deux bords 1 et N :



Montrer que la matrice stochastique P associée à Q et à la mesure π par l'algorithme de Metropolis–Hastings de la première partie est donnée par les valeurs suivantes :

$$P(x, x-1) = \frac{1}{2} \qquad \text{si } x \geq 2;$$
$$P(x, x+1) = \frac{a^{h(x+1)-h(x)}}{2} \qquad \text{si } x \leq N-1;$$
$$P(x, x) = \frac{1 - a^{h(x+1)-h(x)}}{2} \qquad \text{si } 2 \leq x \leq N-1;$$

$P(1,1) = 1 - \frac{a^{h(2)-h(1)}}{2}$ et $P(N,N) = \frac{1}{2}$.

10. Montrer que la plus petite valeur propre λ_N de la matrice P est plus grande que $-a$.

Pour toute paire d'états distincts ($x \neq y$) de $\mathfrak{X}$, on note $c(x,y) = \pi(x)\, P(x,y)$, qui est symétrique en (x,y). On introduit également le chemin $\gamma_{x,y}$ qui est le chemin le plus court reliant x à y dans le graphe de la matrice de transition Q ; si $x < y$, alors $\gamma_{x,y}$ est constitué des transitions $x \to x+1 \to \cdots \to y-1 \to y$, et on a une description similaire si $x > y$. On peut voir $\gamma_{x,y}$ comme une suite d'arêtes orientées $e = (e_- \to e_+)$.

11. Pour $x \neq y$, on pose $D(x,y) = \sum_{f \in \gamma_{x,y}} \frac{1}{(c(f))^{1/2}}$. Montrer la suite d'inégalités suivantes : pour toute fonction $f : \mathfrak{X} \to \mathbb{R}$,

$$\begin{aligned}\mathrm{Var}_\pi(f) &= \frac{1}{2} \sum_{x \neq y} \pi(x)\pi(y)\, (f(x) - f(y))^2 \\ &\leq \frac{1}{2} \sum_{x \neq y} \pi(x)\pi(y)\, D(x,y) \left(\sum_{e \in \gamma_{x,y}} (c(e))^{1/2}\, (f(e_-) - f(e_+))^2 \right) \\ &\leq \frac{1}{2} \sum_{e} c(e)\, (f(e_-) - f(e_+))^2 \left(\frac{1}{(c(e))^{1/2}} \sum_{x,y \,|\, e \in \gamma_{x,y}} \pi(x)\pi(y)\, D(x,y) \right)\end{aligned}$$

où la dernière somme $\sum_e$ est effectuée sur toutes les arêtes orientées $(x \to x+1)$ ou $(x \to x-1)$ du graphe de la matrice Q. En déduire que la seconde valeur propre de P vérifie :

$$\frac{1}{1-\lambda_2} \leq A = \max_e \left(\frac{1}{(c(e))^{1/2}} \sum_{x,y \,|\, e \in \gamma_{x,y}} \pi(x)\pi(y)\, D(x,y) \right).$$

12. Montrer que si $e = (x \to x+1)$ ou $(x+1 \to x)$, alors $c(e) = \frac{\pi(x+1)}{2}$. En déduire que pour tout couple $x < y$,

$$D(x,y) \leq \frac{\sqrt{2}}{(\pi(y))^{1/2}\, (1 - a^{1/2})}.$$

On pourra borner $(\pi(y))^{1/2}\, D(x,y)$ par une série géométrique. Montrer finalement qu'on peut borner supérieurement A par

$$A \leq \frac{2}{(1-a^{1/2})^2}.$$

On pourra remarquer que si $e = (u, u+1)$, alors les x, y tels que $e \in \gamma_{x,y}$ sont $x \in [\![1,u]\!]$ et $y \in [\![u+1, N]\!]$.

13. Déduire des questions précédentes le résultat suivant : le trou spectral ρ de la matrice de Metropolis–Hastings P pour la simulation de la loi π vérifie

$$\rho \geq \rho(a) = \frac{(1-a^{1/2})^2}{2}.$$

14. Modifier légèrement la preuve du théorème C.8 des notes de cours pour établir l'inégalité suivante : si l'on part de l'état initial 1 ($\pi_0 = \delta_1$), alors la loi $\pi_n = \delta_1 P^n$ vérifie :

$$d_{\mathrm{TV}}(\pi_n, \pi) \leq F(a)\, \mathrm{e}^{-\rho(a)\, n}$$

pour tout temps $n \in \mathbb{N}$, et pour une certaine fonction $F(a)$ qu'on explicitera et qui ne dépend que de a (en particulier, elle ne dépend pas de N).

15. Écrire un programme `Python` ou `Sage` qui simule la loi π lorsque $h(x) = x$ et lorsque $h(x) = x + \log x$ (pour tous a et N). Peut-on calculer Z_a pour ces modèles, et en a-t-on besoin ?

Corrigé

1. Les coefficients de la matrice P sont des combinaisons linéaires à coefficients positifs de coefficients de la matrice stochastique Q, donc sont positifs. Sur chaque ligne, on a

$$\begin{aligned}\sum_{y\in\mathfrak{X}} P(x,y) &= P(x,x) + \sum_{y\neq x} P(x,y) \\ &= Q(x,x) + \sum_{y\neq x}(1-A(x,y))\,Q(x,y) + \sum_{y\neq x} A(x,y)\,Q(x,y) \\ &= Q(x,x) + \sum_{y\neq x} Q(x,y) = \sum_{y\in\mathfrak{X}} Q(x,y) = 1,\end{aligned}$$

donc P est bien stochastique.

2. L'hypothèse $0 \leq S(x,y) \leq 1 + \min(T(x,y), T(y,x))$ garantit que les coefficients d'acceptation $A(x,y)$ sont tous compris entre 0 et 1. Vérifions maintenant la réversibilité de la mesure π par rapport à P. Si $x \neq y$, on a

$$\pi(x)\,P(x,y) = \pi(x)\,A(x,y)\,Q(x,y) = \begin{cases} \frac{\pi(x)\,\pi(y)\,Q(x,y)\,Q(y,x)\,S(x,y)}{\pi(x)\,Q(x,y)+\pi(y)\,Q(y,x)} & \text{si } Q(x,y) > 0, \\ 0 & \text{si } Q(x,y) = Q(y,x) = 0. \end{cases}$$

Cette expression est symétrique en x et y, car $S(x,y)$ est une matrice symétrique. On obtient donc le même résultat en calculant $\pi(y)\,P(y,x)$, d'où la réversibilité.

3. Supposons $S(x,y) \in (0,1)$ pour toute paire (x,y) d'états distincts. Alors, le coefficient d'acceptation $A(x,y)$ est toujours compris strictement entre 0 et 1. Comme la matrice Q est irréductible, on en déduit que la matrice P l'est aussi : si x et y sont deux états distincts et $Q(x,x_1)\,Q(x_1,x_2)\cdots Q(x_{n-1},y) > 0$ pour une certaine chaîne d'états $x \to x_1 \to x_2 \to \cdots \to y$, alors on a aussi

$$\begin{aligned}&P(x,x_1)\,P(x_1,x_2)\cdots P(x_{n-1},y) \\ &= (A(x,x_1)\,A(x_1,x_2)\cdots A(x_{n-1},y))(Q(x,x_1)\,Q(x_1,x_2)\cdots Q(x_{n-1},y)) > 0.\end{aligned}$$

Par ailleurs, si x est un état fixé dans $\mathfrak{X}$, on a

$$P(x,x) = Q(x,x) + \sum_{y\neq x}(1-A(x,y))\,Q(x,y) > 0$$

par disjonction des cas : c'est clair si $Q(x,x) > 0$, et sinon, la somme est une combinaison linéaire à coefficients strictement positifs de termes $Q(x,y)$ non tous nuls. Ceci implique l'apériodicité de la chaîne de Markov associée à la matrice P.

Comme on peut choisir arbitrairement les coefficients $S(x,y)$ dans $(0,1)$, on en déduit qu'il existe une infinité (non dénombrable) de matrices P correspondant à ces choix, et d'après la question précédente toutes ces matrices ont pour mesure réversible π.

4. On calcule avec $S(x,y) = 1 + \min(T(x,y), T(y,x))$:

$$A(x,y) = \frac{1+\min(T(x,y),T(y,x))}{1+T(x,y)} = \min\left(1, \frac{1+T(y,x)}{1+T(x,y)}\right) = \min\left(1, \frac{\pi(y)\,Q(y,x)}{\pi(x)\,Q(x,y)}\right)$$

en supposant $Q(x,y) > 0$. Si $Q(x,y) = Q(y,x) = 0$, alors $T(x,y) = T(y,x) = 0$, $S(x,y) = 1$ et $A(x,y) = 1$. La formule pour les coefficients d'acceptation est donc établie.

Supposons π réversible pour Q. Alors, $A(x,y) = 1$ pour tout couple (x,y), donc $P = Q$. Supposons finalement Q irréductible apériodique. Notons que alors si $Q(x,y) > 0$, on a aussi $A(x,y) > 0$, donc :

$$Q(x,y) > 0 \iff P(x,y) > 0.$$

Ceci implique que P est irréductible, et que les chaînes $x \to x_1 \to x_2 \to \cdots \to x_{n-1} \to x$ pour la matrice Q sont les mêmes que pour la matrice P. L'ensemble des périodes de retour $R(x) = \{n \,|\, Q^n(x,x) > 0\}$ est donc le même que $\{n \,|\, P^n(x,x) > 0\}$; et si Q est apériodique, alors il en va de même pour P. Finalement, d'après la question 2., P admet pour mesure réversible π.

5. Développons l'expression donnée pour l'énergie de Dirichlet :

$$\begin{aligned}
&\frac{1}{2}\sum_{x,y\in\mathfrak{X}} \pi(x)\,P(x,y)\,(f(x)-f(y))^2 \\
&= \frac{1}{2}\sum_{x,y\in\mathfrak{X}} \pi(x)\,P(x,y)\,(f(x)^2+f(y)^2) - \sum_{x,y\in\mathfrak{X}} \pi(x)\,f(x)\,P(x,y)\,f(y) \\
&= \frac{1}{2}\sum_{x,y\in\mathfrak{X}} \pi(x)\,P(x,y)\,f(x)^2 + \frac{1}{2}\sum_{x,y\in\mathfrak{X}} \pi(y)\,P(y,x)\,f(y)^2 - \sum_{x,y\in\mathfrak{X}} \pi(x)\,f(x)\,P(x,y)\,f(y) \\
&= \sum_{x,y\in\mathfrak{X}} \pi(x)\,P(x,y)\,f(x)^2 - \sum_{x,y\in\mathfrak{X}} \pi(x)\,f(x)\,P(x,y)\,f(y) \\
&= \sum_{x\in\mathfrak{X}} \pi(x)\,f(x)^2 - \sum_{x\in\mathfrak{X}} \pi(x)\,f(x)\,(Pf)(x) \\
&= \langle (I-P)(f) \,|\, f\rangle_\pi
\end{aligned}$$

en utilisant à la seconde ligne la réversibilité de π par rapport à P. On obtient donc bien la formule souhaitée pour l'énergie de Dirichlet, et comme elle ne dépend que des différences $f(x) - f(y)$, pour toute constante c,

$$\mathscr{E}(f+c) = \mathscr{E}(f).$$

6. Si l'on prend $f_1 = 1$, alors $\pi((f_1)^2) = \sum_{x\in\mathfrak{X}} \pi(x) = 1$, donc f_1 est bien normée ; et

$$(P1)(x) = \sum_{y\in\mathfrak{X}} P(x,y) = 1$$

pour tout $x \in \mathfrak{X}$, donc $P1 = 1$. Ainsi, on peut prendre comme vecteur propre de P pour la valeur propre 1 la fonction constante égale à 1. Alors, les autres fonctions $f_{i\geq 2}$ sont orthogonales à f_1 dans $\mathscr{L}^2(\mathfrak{X},\pi)$, donc

$$0 = \langle f_1 \,|\, f_i\rangle_\pi = \sum_{x\in\mathfrak{X}} \pi(x)\,f_i(x) = \pi(f_i).$$

Ces autres fonctions sont donc de moyenne nulle sous la loi π.

7. Si f est constante à une valeur c, alors sa moyenne $\pi(f)$ vaut également c, donc la variance $\mathrm{Var}_\pi(f) = \pi((f-\pi(f))^2) = \pi((c-c)^2) = \pi(0) = 0$ s'annule. Réciproquement, si $\mathrm{Var}_\pi(f) =$

0, alors f est constante π-presque sûrement, et comme la mesure π charge tout $\mathfrak{X}$, f est une constante.

Montrons maintenant l'identité annoncée pour la seconde valeur propre. Soit f une fonction non constante, qu'on décompose sous la forme $f = \sum_{i=1}^N a_i\, f_i$; l'hypothèse de non constance est équivalente au fait que l'un des coefficients $a_{i\geq 2}$ est non nul. La variance de f s'écrit :

$$\mathrm{Var}_\pi(f) = \pi((f-\pi(f))^2) = \pi\left(\left(\sum_{i=1}^N a_i\,(f_i - \pi(f_i))\right)^2\right) = \pi\left(\left(\sum_{i=2}^N a_i\, f_i\right)^2\right) = \sum_{i=2}^N (a_i)^2$$

en utilisant le caractère orthonormé des fonctions f_i dans l'espace $\mathscr{L}^2(\mathfrak{X},\pi)$. Par ailleurs, l'énergie de Dirichlet étant invariante par translation par les constantes, on obtient :

$$\begin{aligned}\mathscr{E}(f) = \mathscr{E}\left(\sum_{i=2}^N a_i f_i\right) &= \left\langle (I-P)\left(\sum_{i=2}^N a_i f_i\right) \,\middle|\, \sum_{i=2}^N a_i f_i\right\rangle_\pi \\ &= \left\langle \sum_{i=2}^N a_i\,(1-\lambda_i)\, f_i \,\middle|\, \sum_{i=2}^N a_i f_i\right\rangle_\pi = \sum_{i=2}^N (a_i)^2\,(1-\lambda_i).\end{aligned}$$

Le rapport $\frac{\mathscr{E}(f)}{\mathrm{Var}_\pi(f)}$ vaut donc simplement $\frac{\sum_{i=2}^N (a_i)^2\,(1-\lambda_i)}{\sum_{i=2}^N (a_i)^2}$ pour une fonction f non constante. L'identité s'en déduit aisément :

$$\frac{\mathscr{E}(f)}{\mathrm{Var}_\pi(f)} \geq \frac{\sum_{i=2}^N (a_i)^2\,(1-\lambda_2)}{\sum_{i=2}^N (a_i)^2} = 1-\lambda_2$$

pour toute fonction non constante f, avec égalité lorsque $f = f_2$ par exemple.

8. Si x est tel que $|f_i(x)| = \max_{y\in\mathfrak{X}} |f_i(y)|$, alors :

$$\lambda_i\, f_i(x) = (Pf_i)(x) = \sum_{y\in\mathfrak{X}} P(x,y)\, f(y) = P(x,x)\, f_i(x) + \sum_{y\neq x} P(x,y)\, f_i(y);$$

$$|\lambda_i - P(x,x)|\,|f_i(x)| = \left|\sum_{y\neq x} P(x,y)\, f_i(y)\right| \leq \sum_{y\neq x} P(x,y)\,|f_i(x)| = (1-P(x,x))\,|f_i(x)|,$$

d'où $|\lambda_i - P(x,x)| \leq 1 - P(x,x)$. On en déduit que

$$P(x,x) - \lambda_i \leq 1 - P(x,x) \qquad ; \qquad \min_{y\in\mathfrak{X}}(2P(y,y)) \leq 2P(x,x) \leq 1+\lambda_i.$$

C'est en particulier vrai pour $\lambda_i = \lambda_N$ plus petite valeur propre.

9. Les coefficients d'acceptation sont donnés par

$$A(x,x+1) = a^{h(x+1)-h(x)} \qquad ; \qquad A(x,x-1) = 1$$

et $A(x,y) = 1$ pour tout autre couple $x \neq y$. Les formules pour la matrice P s'en déduisent immédiatement.

10. D'après la question 8., $1+\lambda_N \geq 2\ \min_{x\in\mathfrak{X}} P(x,x)$, et le minimum des probabilités de transition $P(x,x)$ est
$$\min_{x\in\mathfrak{X}}\left(\frac{1-a^{h(x+1)-h(x)}}{2}\right) \geq \frac{1-a}{2}.$$
On en déduit que $\lambda_N \geq -a$.

11. La première formule est obtenue de façon analogue à la question 5. Ainsi, si l'on développe la demi-somme des termes $\pi(x)\pi(y)\,(f(x)-f(y))^2$, on obtient :
$$\begin{aligned}&\sum_{x,y\in\mathfrak{X}} \pi(x)\pi(y)\,(f(x))^2 - \sum_{x,y\in\mathfrak{X}} \pi(x)\,\pi(y)\,f(x)\,f(y)\\ &= \sum_{x\in\mathfrak{X}} \pi(x)\,(f(x))^2 - \left(\sum_x \pi(x)\,f(x)\right)^2 = \pi(f^2) - (\pi(f))^2 = \mathrm{Var}_\pi(f).\end{aligned}$$
Pour la seconde formule, remarquons que
$$\begin{aligned}(f(x)-f(y))^2 &= \left(\sum_{e\in\gamma_{x,y}} (f(e_-)-f(e_+))\right)^2 = \left(\sum_{e\in\gamma_{x,y}} \frac{1}{(c(e))^{1/4}}\,(c(e))^{1/4}\,(f(e_-)-f(e_+))\right)^2\\ &\leq \left(\sum_{e\in\gamma_{x,y}} \frac{1}{(c(e))^{1/2}}\right)\left(\sum_{e\in\gamma_{x,y}} (c(e))^{1/2}\,(f(e_-)-f(e_+))^2\right)\end{aligned}$$
par l'inégalité de Cauchy–Schwarz. On reconnaît bien le terme $D(x,y)$, et la seconde ligne est donc établie. Finalement, la troisième ligne s'en déduit en réorganisant les termes de la somme. Bornons chaque terme
$$\left(\frac{1}{(c(e))^{1/2}} \sum_{x,y\,|\,e\in\gamma_{x,y}} \pi(x)\pi(y)\,D(x,y)\right)$$
par la valeur maximale A sur toutes les arêtes $e=(x\to x+1)$ ou $e=(x+1\to x)$. On obtient ainsi :
$$\mathrm{Var}_\pi(f) \leq \frac{A}{2}\sum_e c(e)\,(f(e_-)-f(e_+))^2.$$
Or, la somme de droite est aussi donnée par $\sum_{x,y}\pi(x)\,P(x,y)\,(f(x)-f(y))^2$, donc l'inégalité ci-dessus se réécrit sous la forme :
$$\mathrm{Var}_\pi(f) \leq A\,\mathscr{E}(f)$$
en vertu de la question 5. Autrement dit, pour toute fonction f non constante, $\frac{1}{A} \leq \frac{\mathscr{E}(f)}{\mathrm{Var}_\pi(f)}$. En prenant f qui minimise le rapport énergie de Dirichlet sur variance, on conclut que $1-\lambda_2 \geq \frac{1}{A}$.

12. On calcule facilement
$$c(x,x+1) = \pi(x)\,P(x,x+1) = \frac{1}{Z_a}\,a^{h(x)}\,\frac{a^{h(x+1)-h(x)}}{2} = \frac{1}{2Z_a}\,a^{h(x+1)} = \frac{\pi(x+1)}{2}$$

pour tout $x \in [\![1, N-1]\!]$. Comme π est réversible pour P, on obtient le même résultat avec $c(x+1,x)$. Examinons maintenant pour $x < y$ la quantité $D(x,y)$ (qui est symétrique en x et y). Les valeurs de π étant décroissantes, on peut écrire :

$$\begin{aligned} D(x,y) &= \sum_{t=x}^{y-1} \frac{1}{(c(t,t+1))^{1/2}} = \sum_{u=x+1}^{y} \frac{\sqrt{2}}{(\pi(u))^{1/2}} = \frac{\sqrt{2}}{(\pi(y))^{1/2}} \sum_{u=x+1}^{y} a^{\frac{h(y)-h(u)}{2}} \\ &\leq \frac{\sqrt{2}}{(\pi(y))^{1/2}} \sum_{u=x+1}^{y} a^{\frac{y-u}{2}} \leq \frac{\sqrt{2}}{(\pi(y))^{1/2}\,(1-a^{1/2})} \end{aligned}$$

en bornant la dernière somme par une série géométrique de raison $a^{1/2}$. On en déduit que pour toute arête $e = (u \to u+1)$, on a :

$$\frac{1}{(c(e))^{1/2}} \sum_{x,y \,|\, e \in \gamma_{x,y}} \pi(x)\pi(y)\, D(x,y) \leq \frac{2}{(\pi(u+1))^{1/2}\,(1-a^{1/2})} \sum_{\substack{1 \leq x \leq u \\ u+1 \leq y \leq N}} \pi(x)\,(\pi(y))^{1/2}.$$

La somme sur x peut être bornée par 1, et la somme sur y peut être bornée comme précédemment par $\frac{(\pi(u+1))^{1/2}}{1-a^{1/2}}$. On conclut que pour toute arête e,

$$\frac{1}{(c(e))^{1/2}} \sum_{x,y \,|\, e \in \gamma_{x,y}} \pi(x)\pi(y)\, D(x,y) \leq \frac{2}{(1-a^{1/2})^2},$$

et ceci est donc une borne supérieure sur A.

13. Le trou spectral est la distance minimale que l'on peut mettre en $\{-1,1\}$ et $\{\lambda_2, \lambda_N\}$. Or, on a les inégalités :

$$-1 + (1-a) = -a \leq \lambda_N \leq \lambda_2 \leq 1 - \frac{(1-a^{1/2})^2}{2}.$$

d'après les questions 10. et 12. Le trou spectral vaut donc au minimum

$$\min\left(1-a, \frac{(1-a^{1/2})^2}{2}\right).$$

Pour $a \in (0,1)$, on a trivialement $1-a \geq (1-a^{1/2})^2$, donc le minimum des deux valeurs est $\frac{(1-a^{1/2})^2}{2}$.

14. En modifiant très légérement la preuve de l'inégalité sur la distance en variation totale (dans le cas où la loi initiale π_0 est connue et concentrée en un point x), on obtient

$$d_{\mathrm{TV}}(\delta_x P^n, \pi) \leq \left(\frac{1}{2}\sqrt{\frac{1}{\pi(x)} - 1}\right) \mathrm{e}^{-\rho n}.$$

Ici, $\pi(1) = \frac{a^{h(1)}}{Z_a}$, donc $\frac{1}{\pi(x)} - 1 = \sum_{i=2}^{N} a^{h(i)-h(1)} \leq \frac{a}{1-a}$ en bornant par une somme géométrique. On obtient donc l'inégalité demandée avec $F(a) = \frac{1}{2}\sqrt{\frac{a}{1-a}}$.